Speedchecker Ltd: Recommendations Submitted to FCC in Response to Broadband Deployment Report Notice of Inquiry.

Speedchecker Ltd is pleased to participate in public comment on the FCC Report on the Deployment of Advanced Telecommunications Capability to All Americans in a Reasonable and Timely Fashion.

Since 2008 we have helped millions of users get a better understanding of how to make their Internet go faster. Our solutions empower telecoms, governments and researchers in making their Internet infrastructure better and more available for everyone.

Our active measurement networks and speed checking tools that are used to collect datasets all share the spirit of showing true end-to-end performance as experienced by the end users giving realistic metrics and insights.

Company details:

Speedchecker Ltd, The Black Church, St. Mary's Place, Dublin 7, D07 P4AX, Ireland
Website: http://www.speedchecker.xyz

Author:
Janusz Jezowicz, janusz@speedchecker.xyz

1 Summary

We believe that any benchmark target requires a trusted, accurate and comprehensive measurement methodology. Further, the target benchmark should not only be given as a speed (e.g. 25 Mbps) but also as a Quality of Service requirement.

The agreed benchmarks should be Meaningful, Measurable and Achievable.
Please read the attached file for more information.

Meaningful

A benchmark of 25 Mbps and even 100 Mbps are useful to show how speeds and benchmarks have changed over time but they do little to help determine what this means for the end user. Targets would be better if they included examples of what the end user should expect. For example, should users expect to be able to stream HD or 4K and how much latency or buffering is acceptable.

Measurable

Simple speed test measurements are widely available and produce useful results particularly with regard to seeing if a particular connection meets the required benchmark speed. However, consideration should be given to how to measure the user experience and Quality of Service.

We believe that independent / unbiased crowdsourced internet measurement data have the best opportunity to meet these requirements. Crowdsourced data can be obtained from many independent parties – e.g. private companies such as Speedchecker as well as from research organizations etc.

Achievable

Is it reasonable to expect providers to meet these requirements everywhere or should there be different targets depending on the technology available in the area? Perhaps there should be different benchmark targets depending on the size and remoteness of the population. Different targets allow for the available speeds to be maximised across the country.

This could then be combined with staged targets for the provision of technology in remote areas. E.g. ADSL, then Superfast, Ultrafast and finally Gigabit with timescales longer for Sparse Fringe areas than Cities (Sparse Fringe, Sparse Town, Town, Conurbation then City).

Sample for illustration only	Slow ADSL (10 Mbps)	Fast ADSL (10 Mbps)	Superfast (24 Mbps)	Ultrafast (300 Mbps)	Gigabit (1000 Mbps)
City / Town	Now	Now	Now	5 years	10 years
Conurbation	Now	Now	3 years	10 years	15 years
Town / Village	Now	Now	5 years	12 years	18 years
Sparse Town / Village	Now	5 years	10 years	15 years	20 years
Sparse Fringe	5 years	10 years	15 years	25 years	?

2 The Need for Good Quality Speed Measurement

Regardless of the actual benchmark targets, it is essential that the achieved speeds are measured accurately, independently and consistently and these measurements record the actual user experience.

We believe that crowdsourced internet measurement data have the best opportunity to meet these requirements.

The following points illustrate this argument:

- **Independent / unbiased** – Crowdsourced data can be obtained from many independent parties – e.g. private companies such as Speedchecker as well as from research organizations etc. These are independent measurement companies that can be trusted to provide accurate and trustworthy results. FCC can and should acquire data from many sources and combine them together to ensure that they give unbiased insights as well as ensuring that they complement data sources which may have gaps in coverage.
- **High data granularity** – As stated by NTIA, the FCC Form 477 lacks spatial granularity in certain census blocks which can be larger in rural areas. Crowdsourced internet speed test data are collected with accurate location alongside other collected KPIs. Location is typically collected using GPS or Wi-Fi geolocation which can provide street-level geolocation accuracy. This greatly improves the spatial granularity, particularly in rural areas, to ensure the best possible information.

- **Frequency of updates** – Unlike FCC Form 477 data which is collected twice a year, crowdsourced data are continuously updated and can provide insights much sooner about how the connectivity changes in different areas. This has the added advantage of being able to identify the improved speeds as a result of recent changes.
- **Data recency** – data can be collected in real-time and integrated to mapping.
- **Access type / technology agnostic** – Crowdsourced data can be collected from end user devices such as mobile phones, tablets or computers. The methodology does not limit what access type the user is using e.g. whether it's a fixed or mobile network. The measurements can be made from a user device to a target measurement server on any access type. Depending on the chosen access technology, different KPIs are collected and are available for further analysis. For example, if a user is connected to a fixed broadband via wi-fi, then mobile signal quality data is not available or needed.

  The detailed list of all broadband metrics can be found in Appendix 1

Concerns

Crowdsourced data can raise the following concerns:

- **Data quality** – Crowdsourced data are influenced by many factors such as the user device, network congestions, overloaded testing servers, wi-fi interference and others. Some of these factors can be mitigated easily such as overloaded testing servers or network congestions. Others, such as Wi-Fi interference, need more analysis of the available KPIs (such as wi-fi signal strength) to eliminate results which can influence the data accuracy. In any case, if correct statistical analysis is performed on the large volume of data points the inaccuracies will not affect the end result.
- **Data privacy** – Crowdsourced data contains user location as well as PII data such as IP address. Nevertheless, the PII can be taken out without impacting the analysis options required for broadband mapping. Knowing individual locations of the users is not required as analysis is performed on aggregated data and not individual measurements. Also, time and location data can be altered to a lower precision (e.g. less geographic accuracy and / or the timing of samples can be divided to peak vs off peak and aggregated to monthly samples).
- **Coverage –** While crowdsourced data will not cover every home in the country it can still provide much better data coverage than other alternative approaches such as surveys, or hardware probes (such as Samknows FCC deployment).
- **Increasing load on the network and data costs –** Active measurements (especially repeated tests done in the background) add to the network use and can cause congestion at the network or target measurement server. There needs to be sufficient capacity to perform those measurements otherwise data quality is affected. An alternative approach is to use passive measurement techniques which can collect a lot of KPIs without affecting congestion or consuming precious user mobile data.

Matching crowdsourced data with FCC Form 477 data

To ensure that crowdsourced data can be used to drive a better understanding of broadband availability it's critical to use correct matching methodology between Form 477 data and crowdsourced data. Here are some of the available methods for correct data matching:

- By using the IP address allocation database from ARIN the crowdsourced data can be mapped to specific ISPs automatically.
- GPS/Wi-fi geolocation provides latitude and longitude of the device during the measurement. Samples collected from locations from the census blocks can be aggregated together (after filtering and data sanity checks) and compared with Form 477 data.
- Census blocks with insufficient crowdsourced data should be excluded from analysis and a user recruitment campaign must be targeted to obtain more data.

New approaches

Crowdsourcing data can utilize different approaches:

- **Website data collection** – Utilizing speed test websites such as https://us.broadbandspeedchecker.co.uk can provide a lot of throughput measurements which can show the maximum attainable speed at the time of the test. On the downside, those measurements lack wi-fi or mobile signal data which are crucial for data quality checks.
- **Mobile speed test apps** – Utilizing speed test apps on popular platforms such as iOS or Android can provide lot of measurements as well as a wide range of KPIs (Appendix 1) which provide additional ways to filter data quality.
- **Mobile speed test SDKs** – Rather than relying on a limited set of apps to collect KPIs, the data collection campaign can be extended by implementing SDKs into a lot of different apps which have location permission (as well as user consent). This technique increases the coverage of the data points available for analysis.
- **Active vs passive testing** – Mobile apps can utilize active test methodology which is generating traffic load on the network and measures the KPIs at the time of the increased traffic generation to establish maximum possible speed. Also, there are new alternative approaches which look at existing traffic generated by the user and analyzing network congestion to determine maximum possible speed.

Studies of crowdsourced broadband data

Over the last 10 years there have been studies made on evaluating the use of crowdsourced broadband data for identifying economic benefit to the various communities.

One of the research studies based on Speedchecker data can be found here:

https://www.sciencedirect.com/science/article/pii/S0143622814000782

Other relevant broadband mapping studies include the EU Broadband Mapping project which strives to combine different data points (collected by surveying ISPs and crowdsourced data) onto a single interface and map. More here: https://www.broadbandmapping.eu/

Appendix 1

Field	Description
testId	It's an identifier which uniquely identify the test.
androidDeviceId	It's an identifier which uniquely identify the Android device. AndroidDeviceIDs are created from device's IMEI.
androidFingerprint	AndroidFingerprint describes the information about operating system build and version. The fingerprint can be easily modified by custom versions of Android - it's not standardized. To get the most accurate information about device, it's better to rely on device hardware information fields.
testDate	It contains the date and time the test is performed
clientIpAddress	It's current IP address of the device in the dotted quad format.
downloadKbps	It's a result of the download part of the test, and it's measured in kilobits per second - Kbps.
uploadKbps	It's a result of the upload part of the test and its measured in kilobits per second - Kbps.
latency	It's a result of the latency/ping part of the test and it's measured in milliseconds – ms.
serverName	It's a name of the server which handles the test, usually it's the name of the city where that chosen server is located.
serverCountry	It's a country in which the chosen server is located.
serverCountryCode	It's a code which defines the country location of the chosen server.
clientCountry	The country where the client's device is located.
clientCountryCode	It's a code which defines the country location of the device.
clientCity	The city where the test was proceeded from.
clientLatitude	The latitude of the client's device the test was proceeded from.
clientLongitude	The longitude of the client's device the test was proceeded from.
connectionType	The type of the client's device's connection.

ispName	The ISP name of the client’s device.
networkOperatorName	The network operator name as it was displayed by the client’s device.
networkOperator	It's a combination of mcc (Mobile Country Code) and mnc (Mobile Network Code) of the network registered on the client’s device.

User device information

brand	The brand of the client’s device.
device	The codename of the client’s device created by its manufacturer.
hardware	It's a name of the devices hardware. It's stated by Android kernel.
buildId	The buildID represents the installed version on the client’s device.
manufacturer	The manufacturer of the client’s device.
model	The model of the client’s device.
product	The product code of the client’s device created by its manufacturer.
locationType	The method used for determination of the client’s device location.
simNetworkOperatorName	The operators name of the sim card installed in the client’s device.
simNetworkOperator	Almost the same as *networkOperator*, but it's operator which is associated with the sim card, not the network which is registered the client’s device. There are of course cases when both operators are the same so then *simNetworkOperator* is null. .

connectionType	The type of the client's device's connection (e.g. GPRS,3G,4G, wi-fi)
testType	testType defines the communication protocol used by the test. The test makes a direct TCP connection using Websocket protocol with the server used for the test whenever its possible. In cases in which it's not possible, the http fallback mechanism is used. (1: http, 2: https, 3: ws, 4: wss)

Wi-Fi Network information

WifiNetworkAuth	*Authentification type of the clients wifi network.*
WifiNetworkChannel	*Channel number of the clients wifi network.*
WifiNetworkRouterBrand	*The brand of the clients wifi router.*
WifiNetworkFreq	The frequency of the clients wifi network.
WifiNetworkSignalStrengh	The strength of the signal to the wifi base station from the clients device.
WifiConflictingNetworks	The number of networks which are on the same channel.
WifiNeighouringNetworks	The number of neighbouring networks with the clients network.

Mobile Network information

signalCellType	signalCellType reflects the network connection type associated with the signal and cell information provided by the device. 1 = GSM, 2 = CDMA, 3 = WCDMA, 4 = LTE
mcc	A three digit code which defines the country where the network operator is located - Mobile Country Code.

mnc	A three digit code which specifies the network operator in particular country - Mobile Network Code. When its combined with mcc - Mobile Country Code - then every mobile network can be uniquely identified.
pci	LTE Physical Cell Identity. An integer to identify the physical LTE cell the user is connected to. The value is unique to the physical cell antennae rather than a specific cell tower. Valid values are 0 to 503. A value of 65535 or null indicates that the device was unable to return a PCI value.
tac	LTE Tracking Area Code. A 16 bit integer used to facilitate handoff of a device between cells. The Tracking Area Identity can be determined by prepending the MCC and MNC to the Tracking Area Code.
baseStationId	Base Station Id 0..65535, Integer. MAX_VALUE if unknown
baseStationLatitude	cdma base station latitude in units of 0.25 seconds, Integer.MAX_VALUE if unknown
baseStationLongitude	cdma base station longitude in units of 0.25 seconds, Integer.MAX_VALUE if unknown
networkId	cdma network identification number, -1 if unknown
systemId	cdma system identification number, -1 if unknown
cid	gsm cell id, -1 if unknown, 0xffff max legal value
lac	gsm location area code, -1 if unknown, 0xffff max legal value
psc	primary scrambling code for UMTS, -1 if unknown or GSM
asuLevel	signal level as an asu value between 0..31, 99 is unknown Asu is calculated based on 3GPP RSRP. Refer to 3GPP 27.007 (Ver 10.3.0) Sec 8.69 .
dbm	signal strength as dBm

level	signal level as an int from 0..4
timingAdvance	the timing advance value for LTE, as a value in range of 0..1282. Integer.MAX_VALUE is reported when there is no active RRC connection. Refer to 3GPP 36.213 Sec 4.2.3